



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

A single nucleotide difference at the 3' end of an intron causes differential splicing of two histocompatibility genes

Citation for published version:

Archibald, AL, Thompson, NA & Kvist, S 1986, 'A single nucleotide difference at the 3' end of an intron causes differential splicing of two histocompatibility genes', *EMBO Journal*, vol. 5, no. 5, pp. 957-65.

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Publisher's PDF, also known as Version of record

Published In:

EMBO Journal

Publisher Rights Statement:

IRL
Press
Limited,
Oxford,
Englan

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



A single nucleotide difference at the 3' end of an intron causes differential splicing of two histocompatibility genes

Alan L. Archibald¹, Nancy A. Thompson and Sune Kvist

Swiss Institute for Experimental Cancer Research, Chemin des Boveresses, CH-1066 Epalinges s/Lausanne, Switzerland

¹Present address: A.F.R.C. Animal Breeding Research Organization, West Mains Road, Edinburgh EH9 3JQ, UK

Communicated by B. Dobberstein

The murine histocompatibility class I genes, H-2 K^b and K^k, display considerable homology at their 3' ends. In fact, from exon 5 to the termination codon, only two nucleotides differ between the two genes, one at the 5' end and the other at the 3' end of intron 7. Despite this similarity, the gene products have distinctly different mol. wts as determined by SDS-PAGE. By constructing two hybrid genes, pC2 and pC4, we demonstrated that it is the cytoplasmic parts of the antigens (encoded by exons 6–8) which are responsible for the major difference in mol. wt. We have used site-directed mutagenesis to change the two nucleotides in intron 7 of the H-2 K^k gene to those present in the H-2 K^b gene. S1 nuclease mapping has been used to identify the actual splice site of the authentic K^b and K^k genes, the hybrid genes and the mutagenized genes. We have shown that it is the 3' nucleotide difference, nine nucleotides upstream of the 3' splice site, which causes the different excision of intron 7 of the K^b gene. The 5' nucleotide difference does not alter the splicing. The choice of branch points and 3' splice signals for intron 7 of five H-2 class I genes, is discussed.

Key words: histocompatibility genes/intron acceptor sites/RNA splicing/site-directed mutagenesis/S1 mapping

Introduction

Most genes in higher eukaryotes have their protein-coding regions (exons) interrupted by non-coding regions of DNA (introns). The primary transcripts (pre-mRNAs) of structural genes thus contain intervening sequences, which are removed to produce the mature mRNA. Our knowledge of splicing of pre-mRNAs has recently advanced considerably. This is mainly the result of the development of *in vitro* splicing systems (Green *et al.*, 1983; Padgett *et al.*, 1983; Hernandez and Keller, 1983; Krainer *et al.*, 1984; Grabowski *et al.*, 1984) and the characterization of the intermediates from both *in vitro* (Ruskin *et al.*, 1984; Padgett *et al.*, 1984) and *in vivo* (Pikielny *et al.*, 1983; Zeitlin and Efstratiadis, 1984; Rodriguez *et al.*, 1984; Domdey *et al.*, 1984) splicing reactions. Analysis of the reaction products has established a two-stage pathway for the removal of the intron. First, the pre-mRNA is cleaved at the 5' splice site to produce the first exon. Secondly, the intron is excised by cutting its 3' end with concomitant ligation of exon 1 to exon 2. The excised intron as well as the intermediate intron–exon 2, have the unusual form of a lariat (for reviews, see Weissman, 1984; Keller, 1984; Sharp, 1985). This lariat structure has been shown to exist during the splicing of yeast pre-mRNAs (Rodriguez *et al.*, 1984; Domdey *et al.*, 1984), an adenovirus-2 major late transcript (Padgett *et*

al., 1984) and β -globin RNA precursors (Ruskin *et al.*, 1984; Zeitlin and Efstratiadis, 1984). Thus, the lariat configuration seems to be a common feature of the splicing reaction in both yeast and higher eukaryotes. It is therefore reasonable to assume that the formation of the lariat structure plays a crucial role in the maturation of mRNA.

Analysis of the exon–intron boundaries has revealed that introns begin with the 5' nucleotides GT and end with the 3' nucleotides AG (Breathnach *et al.*, 1978). After the initial cleavage of the intron at the 5' end, the guanosine residue in the 5' GT attacks an adenosine residue close to the 3' end of the intron to form a 2'–5' phosphodiester bond, called the branch site (Wallace and Edmonds, 1983). The adenosine nucleotide involved in the branch site has been identified for several introns and has been found to be located 18–37 nucleotides upstream from the AG dinucleotide ending the intron (Konarska *et al.*, 1985; Reed and Maniatis, 1985; Ruskin *et al.*, 1985). Inactivation of the branch point nucleotide by deletion does not prevent splicing but activates alternative sites (cryptic branch sites) located close to the authentic branch point (Reed and Maniatis, 1985; Ruskin *et al.*, 1985). In an extensive study of the large intron of the rabbit β -globin gene, it was shown that a minimum intron length of 80 nucleotides was necessary for efficient and correct splicing (Wieringa *et al.*, 1984). In addition, it was found that only the first six nucleotides at the 5' end and the 24 nucleotides at the 3' end were required.

In all introns so far examined, the region proximal to the 3' splice site is rich in pyrimidines (C+T) and this conservation might indicate an important role for some of these nucleotides (Wieringa *et al.*, 1984; Frendewey and Keller, 1985; Ruskin and Green, 1985). The adenosine branch nucleotide is located just upstream of the pyrimidine-rich stretch of nucleotides and in yeast a conserved sequence 5'-TACTAAC-3' is found 20–55 nucleotides from the 3' end of introns (Langford *et al.*, 1984). The adenosine closest to the 3' end of this heptanucleotide is the residue involved in the branch formation. It is not clear whether base pairing interactions are involved between the 5' splice site and the TACTAAC box. In higher eukaryotes a more general consensus sequence PyTPuAPy has been noted just upstream of the pyrimidine tract (Keller and Noon, 1984; Ruskin *et al.*, 1984). Limited information is available about the nucleotide composition of the 3' pyrimidine tract required for efficient and correct splicing.

For a mouse histocompatibility class I gene alternative splicing seems to occur at intron 7 separating exons 7 and 8 (Kress *et al.*, 1983). Mouse H-2 class I genes are located in the major histocompatibility complex (MHC) and encode the H-2 K, D and L antigens. The function of the histocompatibility antigens is to restrict cytolytic T cells to lyse only infected cells (Zinkernagel and Doherty, 1979). These antigens consist of a heavy chain (mol. wt 43 000) non-covalently associated with β_2 -microglobulin (mol. wt 12 000). The heavy chain is an integral membrane protein with the largest part of the protein exposed on the cell surface and the remaining one-fourth of the protein spanning the plasma membrane. About 30 amino acids protrude on the cytoplasmic

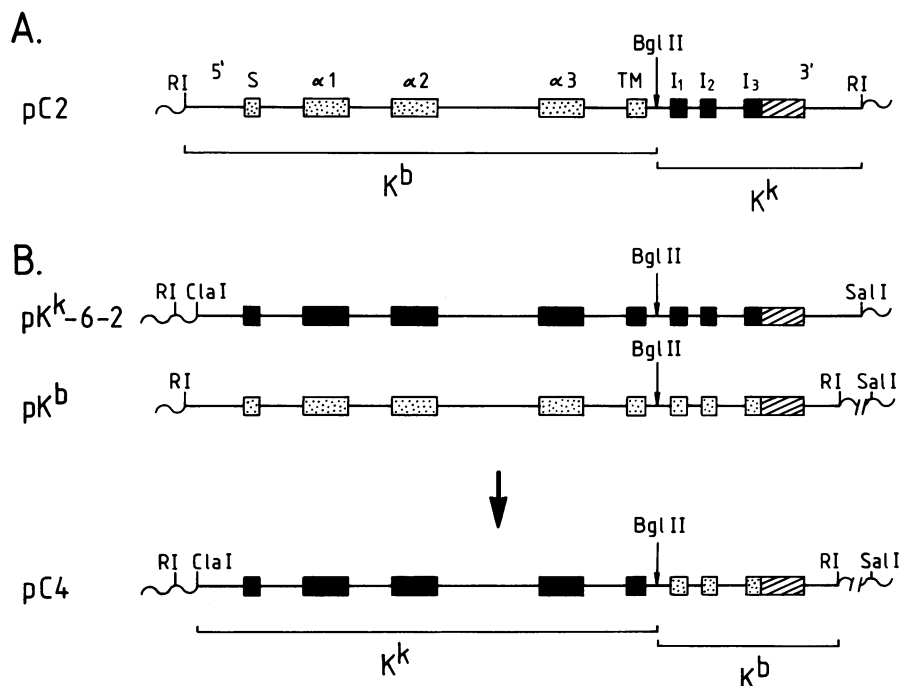


Fig. 1. Structure of the hybrid H-2 K genes pC2 and pC4. The hybrid genes were constructed by using the *Bgl*II site in intron 5. The pC2 gene has the extracellular exons (1–5) of H-2 K^b origin and the exons encoding the cytoplasmic part (6–8) of H-2 K^k origin (A), whereas the pC4 gene represents the converse arrangement (B). Exons are shown as boxes, either solid (H-2 K^k) or dotted (H-2 K^b). The 3'-untranslated regions are shown as hatched boxes. Vector sequences (pBR322) are denoted by wavy lines. Restriction enzymes are indicated, RI being *Eco*RI. The mutated genes m77 and m80 (see text) were constructed in a similar way to pC4. The 3' *Bgl*II/*Sal*I fragment was derived from the subclone pK^k-16-5. See Materials and methods for details.

side of the membrane and, interestingly, this part of the protein is encoded by three separate mini-exons (exons 6, 7 and 8; for a review, see Steinmetz and Hood, 1983).

We have recently determined the DNA sequence for the H-2 K^k gene (Arnold *et al.*, 1984b) and compared it with that of the H-2 K^b gene (Weiss *et al.*, 1983). Here we demonstrate that their pre-mRNAs are differently spliced at intron 7 and that a single nucleotide difference at the 3' end of intron 7 is responsible for this difference.

Results

We have previously constructed a recombinant gene of H-2 K^b and K^k origin (Arnold *et al.*, 1984a). This gene (called pC2) has exons 1–5 from the K^b gene and exons 6–8 from the K^k gene (Figure 1A). The site of recombination is the endogenous *Bgl*II restriction site located in the intron between exons 5 and 6. Therefore, the C2 antigen, encoded by the pC2 hybrid gene, has the extracellular part of the molecule, as well as the membrane-spanning segment, from the H-2 K^b antigen, whereas the cytoplasmic region is of H-2 K^k origin. Close inspection of the DNA sequences of the K^b and K^k genes reveals that from the *Bgl*II site (site of recombination) through to the 3'-non-coding region, only two nucleotide differences are present. No differences are found in any of the cytoplasmic exons (exons 6, 7 and 8). One would, therefore, expect the C2 antigen to have a mol. wt identical to that of the H-2 K^b antigen.

We have transfected 1T 22-6 cells (a mouse fibroblast cell line of H-2^a origin) with the pC2 hybrid gene as well as with the parental genes (H-2 K^b and K^k). Correct cell surface expression of the corresponding antigens was verified by fluorescence-activated cell sorter (FACS) analysis (Arnold *et al.*, 1984a). To determine the mol. wt of the C2 antigen we pulse-labeled the

cells with [³⁵S]methionine for 20 min, lysed the cells and immunoprecipitated the antigens with monoclonal antibodies. The precipitates were analysed by SDS-polyacrylamide gel electrophoresis (SDS-PAGE). Surprisingly, the C2 antigen has a mol. wt clearly distinct from the K^b antigen and is also slightly smaller than the H-2 K^k antigen (Figure 2). This finding could be explained in at least two ways: (i) the mRNA for the C2 antigen is spliced differently at exons 6–8 when compared with the K^b gene, giving rise to an antigen of a lower mol. wt; (ii) the C2 antigen is post-translationally modified (for example, by proteolytic cleavage or modification of carbohydrate groups). We found the first explanation more likely and decided to test it experimentally.

To examine whether or not the cytoplasmic exons were responsible for the difference in mol. wt observed between the C2 and K^b antigens we constructed the pC4 gene (Figure 1B and Materials and methods). This hybrid gene has exons 1–5 of H-2 K^k origin and exons 6–8 of K^b origin. The pC4 gene was introduced into 1T 22-6 cells, along with the neophosphotransferase gene and G-418-resistant clones expressing the C4 antigen were identified by FACS analysis using monoclonal antibodies (not shown). Several of these cell clones were labeled with [³⁵S]-methionine and analysed by SDS-PAGE as described above. The C4 antigen has a mol. wt very similar or identical to that of the K^b antigen (Figure 3, cf. lanes 2 and 3), and clearly distinct from that of the K^k antigen (cf. lanes 1 and 2). From these experiments, we conclude that the exons encoding the cytoplasmic part of the H-2 K^b and K^k antigens are responsible for the observed difference in mol. wt.

Site-directed mutagenesis of intron 7 of the H-2 K^k gene

As no cDNA clone is available for the H-2 K^k antigen, the precise structure of the C terminus is not known. The H-2 K^b

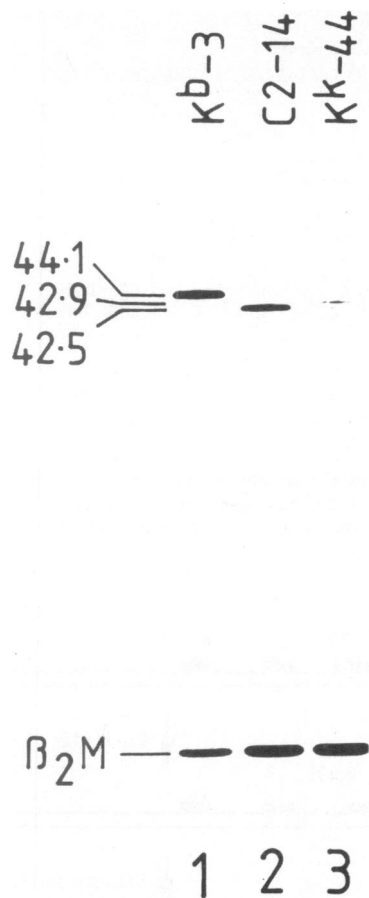


Fig. 2. SDS-PAGE analysis of IT 22.6 cells expressing H-2 K^k , K^b and C2 antigens. Immunoprecipitation was carried out with antibody 28-8-6s (anti- K^b , lanes 1 and 2) and antibody H100-27.55 (anti- K^k , lane 3). The mol. wts of the heavy chain are indicated in kilodaltons. The position of the light chain (β_2M) is also shown. The cell clones used are given at the top.

mRNA encodes 10 amino acids in the last exon (Reyes *et al.*, 1982), whereas the H-2 D^d mRNA has only one amino acid (Brégère *et al.*, 1981; Sher *et al.*, 1985). Comparison of the DNA sequences of the K^b and K^k genes reveals only two nucleotide differences, one at the 5' end of intron 7 (K^b GTG to K^k GTA) and another in the 3' pyrimidine tract of the same intron (K^b GTC to K^k GAC; Figure 4). Two further differences in the middle of intron 7 (K^b CAGG-TT; K^k CATGATT) have been reported (Weiss *et al.*, 1983). However, we find that the K^b and K^k sequences are identical in this region. If alternative splicing is responsible for the differences found in mol. wt between the K^b and K^k antigens, it is likely that this is due to the nucleotide differences in intron 7. We decided to analyse this by changing the two nucleotides in intron 7 of the H-2 K^k gene, separately, to those present at the identical positions in the K^b gene in order to see how this would affect: (i) the mol. wt of the antigen and (ii) the mature mRNA.

To change each of these two nucleotides in intron 7 of the K^k gene, a subclone of the K^k gene was nicked specifically at the *Clal* site with the *Clal* enzyme in the presence of ethidium bromide. The nicked DNA was then digested with exonuclease III in a controlled manner to produce partially single-stranded DNA. A synthesized oligonucleotide containing the relevant mismatch (corresponding to the sequence of the K^b gene; K^k 4606 and K^k 4707 for the 5' and 3' nucleotides, respectively) was then annealed and polymerization was completed with the Klenow

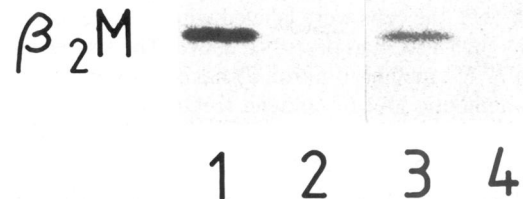
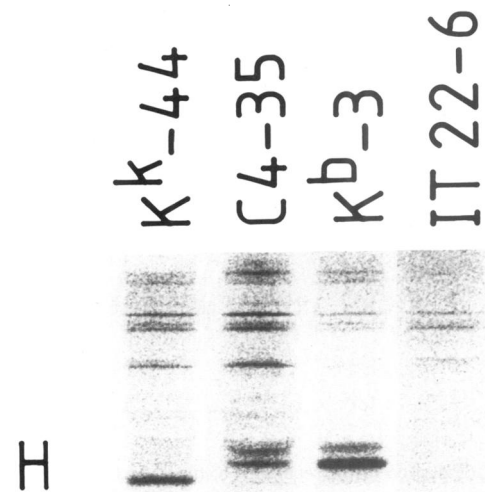


Fig. 3. SDS-PAGE analysis of IT 22-6 cells expressing H-2 K^k , K^b and C4 antigens. Immunoprecipitation was carried out with antibody 28-8-6s (anti- K^b , lanes 3 and 4) and antibody H100-27.55 (anti- K^k , lanes 1, 2 and 4). The positions of the heavy chain (H) and the light chain (β_2M) are shown. The cell clones used are given at the top.

fragment of the DNA polymerase (Figure 5; see Materials and methods for details). After ligation and transfection of bacteria with the DNA, colonies were screened with the ^{32}P -labeled oligonucleotides. Colonies containing mutated plasmids were detected by differential hybridization to the 'mutagenic' oligonucleotide (Wallace *et al.*, 1981).

Twenty two positive colonies were selected after the first screening for each mutagenesis. Plasmid DNA was prepared from these clones and used to transform *Escherichia coli* 71/18 cells. The transformed bacteria were screened with the oligonucleotides as before. This step was necessary in order to ensure that colonies identified as positive during the first round screen were not of mixed genotype, i.e. contain wild-type and mutant plasmids. Of the 44 colonies picked from the positives during the

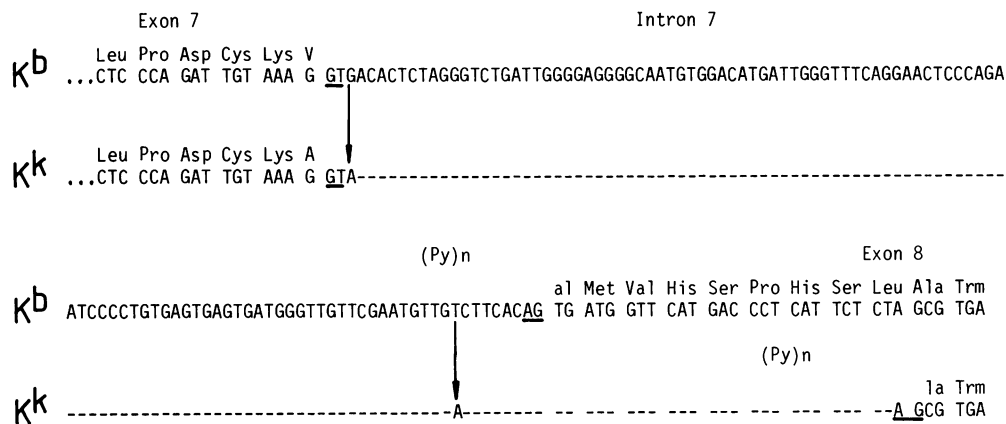


Fig. 4. Comparison of the sequences of the 3' end of the H-2 K^k and K^b genes. The sequences for exon 8 and the 3' end of exon 7 are given in full for both genes along with the corresponding amino acids. The nucleotide differences in intron 7 are arrowed. For clarity nucleotides in intron 7 of K^k which are identical to those present in K^b are shown as dashes (-) except for key sequence features which are given in full for both genes. Acceptor and donor splice sites are underlined. The pyrimidine-rich tracts are denoted (Py)n.

first round screen, 20 were found to be of mixed genotype.

Plasmid DNA was prepared from positive colonies and the DNA sequence was determined for intron 7. One of these clones, K^k -16-5-m77, has the 5' nucleotide of intron 7 mutated to the guanosine present at the corresponding position in the H-2 K^b gene. Another clone, K^k -16-5-m80, carries the thymidine residue of the K^b gene at the 3' end of intron 7. These two clones were used to reconstruct two new H-2 K^k genes (m77 and m80) which differ only by a single nucleotide in intron 7. The construction of m77 and m80 was done in a similar way to pC4 (see Figure 1B and Materials and methods for details).

1T 22-6 cells were transfected with the genes m77 and m80, respectively and cell clones expressing the corresponding antigens were identified by FACS analysis as before. To determine whether or not the mutagenized genes encoded proteins with different mol. wts, the cells were labeled with [35 S]methionine and analysed by SDS-PAGE as described above. The antigen encoded by gene m77 (5' mutant in intron 7) has exactly the same mol. wt as the authentic H-2 K^k antigen (Figure 6, cf. lanes 3 and 4). In contrast, the antigen encoded by the m80 gene co-migrates with the H-2 K^b antigen (Figure 6, cf. lanes 1 and 2). The antibody did not precipitate anything from the non-transfected 1T 22-6 cells (lane 5). The 5' mutation gene (m77) does not cause any size shift, whereas the 3' mutation (gene m80) changes the apparent mol. wt from that of the H-2 K^k antigen to that of H-2 K^b . Thus, it seems that a single base in the 3' region of intron 7 of the H-2 K^k gene is crucial for the splicing pattern of this intron. To prove this, and to precisely map the splicing point of the H-2 K^k gene, as well as of the mutants, we performed the S1 nuclease digestion experiment described below.

S1 nuclease mapping of the 3' end of the H-2 K genes

mRNA was isolated from cells expressing the mutant H-2 K^k genes (m77 and m80), the hybrid genes pC2 and pC4, as well as the parental K^k and K^b genes (see Materials and methods for details). To protect the relevant fragments of mRNA from digestion with the S1 nuclease, we used a DNA probe spanning almost the entire 3' end of the gene. This probe originated from a subclone of the H-2 K^k gene and extends from the middle of exon 5 (membrane exon) to the 3'-non-coding region (Figure 7A). Thus, it is 776 nucleotides long and contains exons 5–8 as well as introns of the 3' end of the K^k gene, including 164 nucleotides

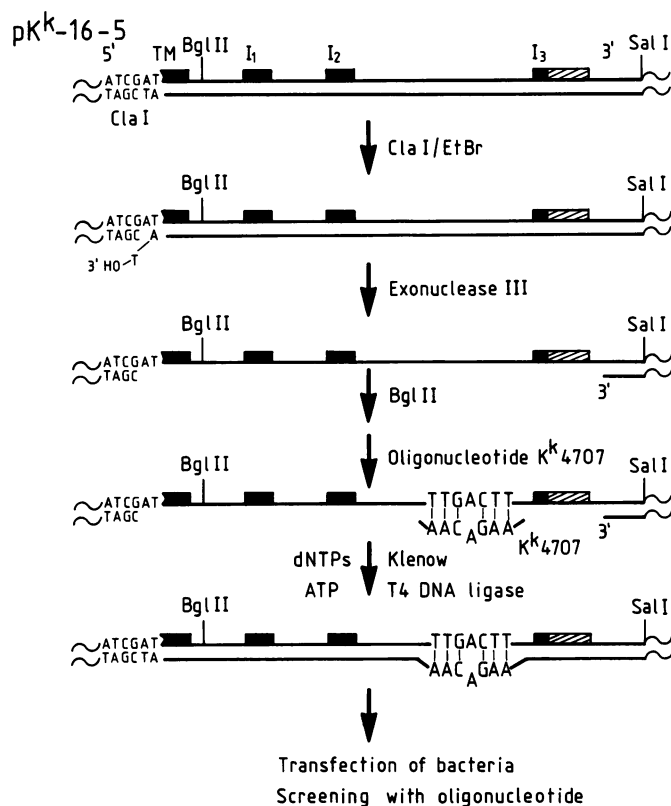


Fig. 5. Strategy for the oligonucleotide site-directed mutagenesis. The subclone p K^k -16-5 was nicked with *Cla*I in the presence of ethidium bromide (EtBr) to produce a free 3'-hydroxyl end. The plasmids were then rendered partially single stranded by digestion with exonuclease III. Specific oligonucleotides were annealed and the plasmids were made double stranded by the Klenow polymerase and closed by T4 DNA ligase. Bacteria transfected with the manipulated plasmids were screened with appropriate 32 P-labeled oligonucleotides. The plasmids shown do not represent the total spectrum of molecules generated during the experiment. Following the nicking with *Cla*I, some plasmids will be nicked on the sense strand, whereas others will be linearized or completely unaffected. The *Bgl*III digestion should linearize all the unwanted plasmids, including those single stranded in the vector sequence upstream of the *Cla*I site. The desired form shown in the figure should be resistant to *Bgl*III digestion. Enzymes used are indicated. Vector (pBR322) and H-2 K gene sequences are shown as wavy and straight lines, respectively. For further details see text and Materials and methods.

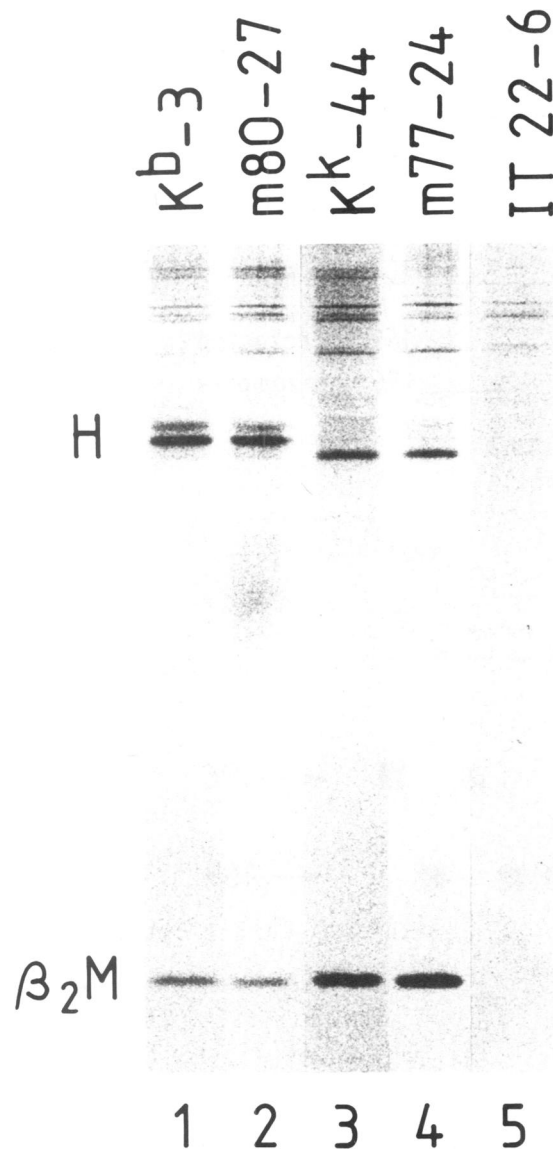


Fig. 6. SDS-PAGE analysis of IT 22-6 cells expressing H-2 K^b, K^k m77 and m80 antigens. Immunoprecipitation was carried out with antibody 28-8-6s (anti-K^b, lanes 1 and 5) and antibody H100-27.55 (anti-K^k, lanes 2, 3, 4 and 5). The positions of the heavy chain (H) and the light chain (β_2 M) are shown. The cell clones used are given at the top.

of the 3'-untranslated region. This fragment was ligated into the *Eco*RI–*Hinc*II sites of M13mp8 and a uniformly labelled minus strand DNA probe was synthesized by using the pentadecamer M13 primer (see Materials and methods for details).

The labeled probe was hybridized to mRNA, digested with S1 nuclease and protected fragments were analysed by polyacrylamide gel electrophoresis. Only one of the protected mRNA fragments varied from one cell line to the other. This band corresponds to a fragment of either 169 or 196 nucleotides, and must therefore contain part of the 3'-untranslated region. Exon 8 contains both coding information and the untranslated region and, thus, must represent the fragment which varies in length due to alternative splicing at the 5' end of this exon. A sequence determination of the probe was made and run in parallel on the same gel. This made possible an accurate measurement of the length of the different fragments. The H-2 K^b-specific band differs by 27 nucleotides from that of the H-2 K^k (Figure 7B, lanes 2 and

3, respectively). The K^b mRNA is known to have 10 amino acids encoded by exon 8 (Reyes *et al.*, 1982) whereas the K^k mRNA must use an acceptor site 27 nucleotides downstream. This leaves only two coding nucleotides before the termination codon. Our construction pC2 which has the 3' part of the gene of K^k origin is spliced identically to the K^k mRNA, whereas pC4 mRNA is spliced in the same way as the K^b mRNA (Figure 7B, lanes 4 and 5, respectively). Most interestingly, the mRNA of gene m77 (5' mutant of intron 7) is spliced in exactly the same way as the K^k mRNA, whereas the band for the m80 mRNA (3' mutant of intron 7) is shifted to the position corresponding to that of K^b (Figure 7B, lanes 6 and 7, respectively). Histocompatibility genes constitute a multigene family which includes > 30 genes. Therefore, we included mRNA from the parental IT 22-6 cell line to account for the background (Figure 7B, lane 1). Fragments corresponding to exons 6 and 7 could also be visualized on a high percentage gel. No variation was seen in any of these bands indicating that these exons are spliced identically for both H-2 K^b and K^k mRNAs (data not shown). We conclude that our original observation of different mol. wts for the K^b and C2 antigens is due to differentially spliced mRNA encoding the last exon (exon 8) of the antigen.

Discussion

The DNA sequences of two mouse class I histocompatibility genes (H-2 K^b and K^k) show considerable homology at their 3' ends. In fact, only two nucleotides are different in the last intron, one at the 5' end and the other at the 3' end (Figure 4). Despite this homology, a hybrid K^b and K^k gene, pC2, encoded an antigen with a mol. wt which indicated differences in the mRNA splicing at the 3' end. We have confirmed this by three independent methods: (i) construction of the hybrid gene pC4; (ii) site-directed mutagenesis in intron 7 of the H-2 K^k gene; (iii) S1 nuclease mapping of the parental, constructed and mutated genes.

The C4 hybrid antigen is K^k in the extracellular and the membrane-spanning segment of the molecule, but is K^b in the cytoplasmic part. This molecule has a mol. wt identical to the H-2 K^b antigen which is different from the H-2 K^k antigen (Figure 3).

We have used site-directed mutagenesis to change separately the two nucleotides in intron 7 of the H-2 K^k gene to the corresponding ones of the H-2 K^b gene (Figure 5). The mutated genes were introduced into IT 22-6 cells. The gene mutated at the 5' end of intron 7 (m77) encoded an antigen with an apparent mol. wt identical to that of the parental H-2 K^k antigen, whereas the 3' mutation (m80) changed the mol. wt of the antigen to that of the H-2 K^b antigen (Figure 6). By using S1 nuclease mapping we could demonstrate that the m80 mRNA is spliced identically to the H-2 K^b and pC4 genes. All three genes have 10 amino acids in exon 8 whereas the m77 gene and the H-2 K^k gene are spliced to give only a single amino acid in exon 8 (Figure 7B). We draw the following conclusions: (i) the mol. wt of the C4 antigen proves that the cytoplasmic part of the antigen influences the apparent mol. wt; (ii) the nucleotide difference at the 5' end of intron 7 of the K^b and K^k genes influences neither the mol. wt of the K^k antigen (m77 is identical to K^k) nor the mRNA splicing; (iii) the 3' nucleotide difference in intron 7 alters the splicing of the mRNA and thus the mol. wt of the antigen; (iv) the authentic H-2 K^k mRNA is spliced to give only a single amino acid encoded by exon 8. Thus, a single nucleotide, in the 3' region of an intron, can change the splicing so that an alternative acceptor site is used.

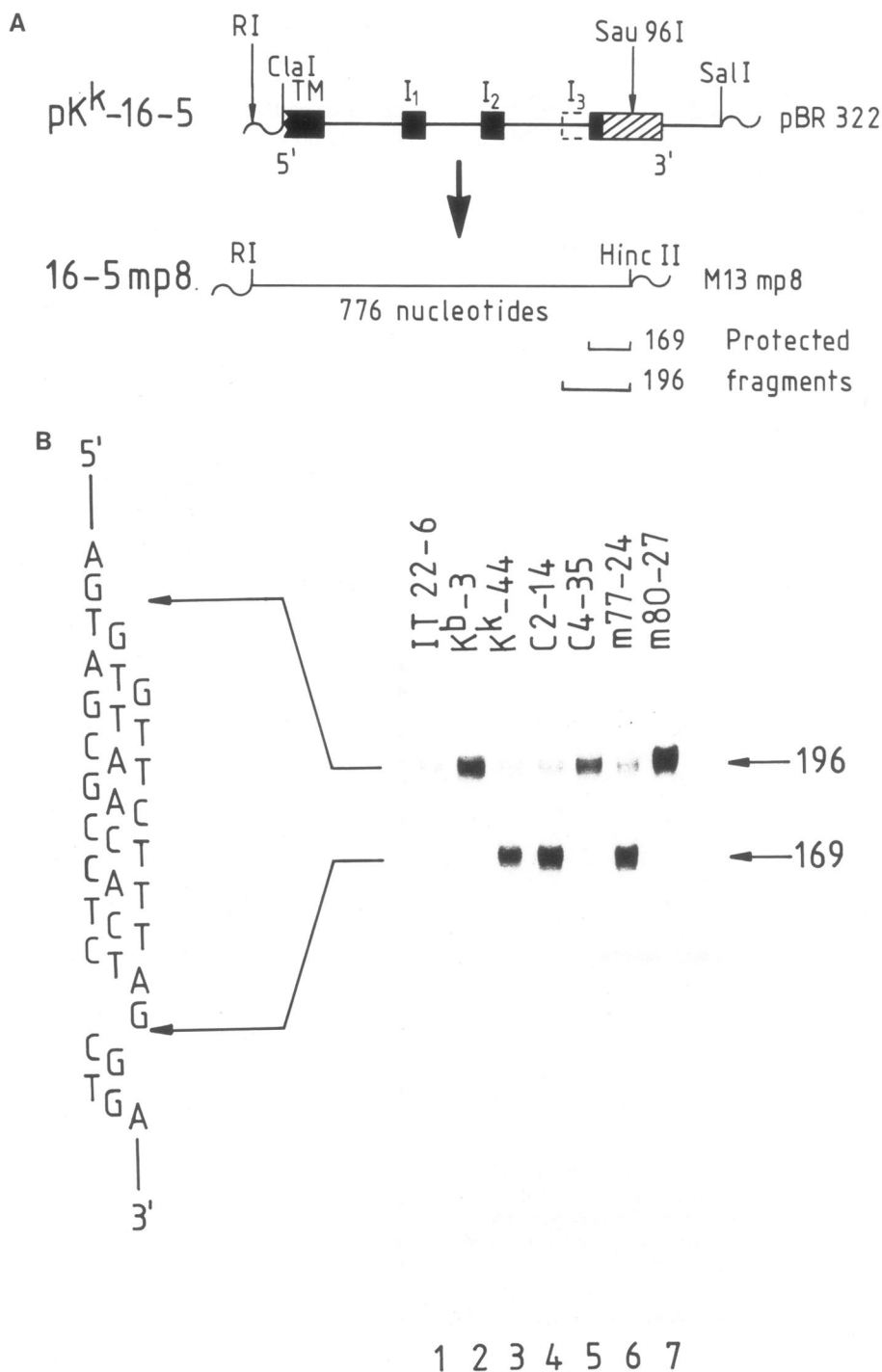


Fig. 7. S1 nuclease analysis of the 3' splicing sites of H-2 K genes. (A) A fragment of 776 nucleotides, spanning the entire 3' end of the K^k gene, was isolated by *EcoRI*/*Sau96I* digestion of clone pK^k-16-5 DNA. This fragment was made blunt at the *Sau96I* site and inserted into the *EcoRI*/*HincII* sites of an M13mp8 vector. A ³²P-labeled, single-stranded probe was synthesized and annealed to mRNA. Filled boxes denote exons 5–8; TM is the exon encoding the transmembrane domain, I₁, I₂ and I₃ are the cytoplasmic exons and the hatched box indicates the 3'-untranslated region. The variable size of exon 8 is shown as a dotted box. The expected size of the fragments protected by the probe against S1 nuclease digestion are denoted for exon 8 only. Restriction enzyme sites are indicated, RI being *EcoRI*. (B) mRNA was isolated from cell clones, indicated in the figure, as described under Materials and methods. Radioactively labelled probe was annealed to mRNA, digested with S1 nuclease and the protected fragments were analysed by polyacrylamide gel electrophoresis. Positions of the 196- and 169-nucleotide bands were identified by co-running of the pK^k-16-5 fragment which had undergone the dideoxy sequencing reactions. The distance between the bands is 27 nucleotides. The DNA sequence shows the alternative splice sites and the region between these sites.

It should be noted that the 3' nucleotide difference is a thymidine residue in the K^b gene and an adenosine in the H-2 K^k gene (Figure 4). One of the intermediate products during splicing is a lariat structure involving a 2'–5'-phosphodiester bond

formed between the guanosine at the 5' end of the intron and an adenosine residue at the 3' end (Wallace and Edmonds, 1983; Padgett *et al.*, 1984; Ruskin *et al.*, 1984; Zeitlin and Efstratiadis, 1984). The sequences flanking the K^k adenosine residue

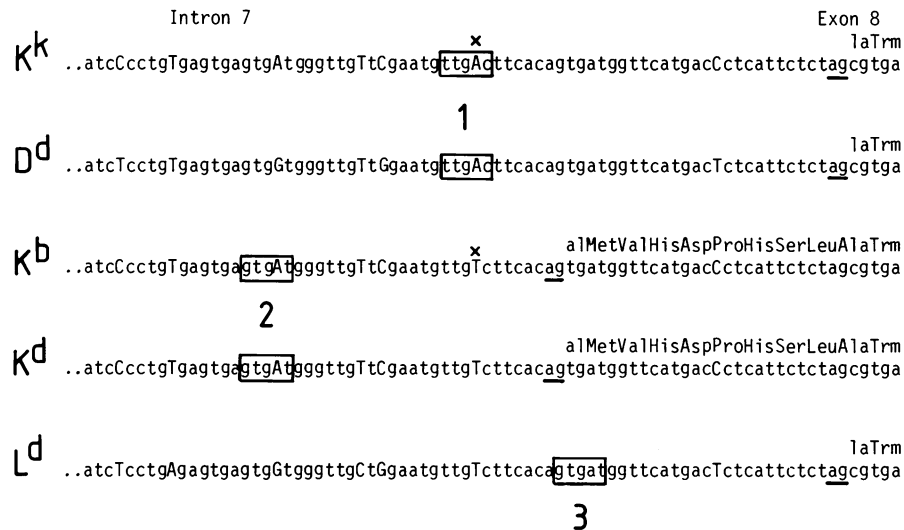


Fig. 8. Comparison of the sequences at the 3' end of five H-2 class I genes. Nucleotides common to all five genes are shown in lower case letters. Nucleotides which differ amongst the genes are shown in capitals. The amino acid translation of exon 8 for these five genes is given. The boxes denote the putative branch sites for each of the five genes. The sites have been numbered 1–3 (see text for explanation). The acceptor splice sites are underlined. The nucleotides involved in the K^k to K^b mutagenesis experiment are marked with crosses (x).

(TTCAG), conform to the consensus sequence for a branch site (PyTPuAPy) as proposed by Keller and Noon (1984) and Ruskin *et al.* (1984). Furthermore, the adenosine would be the nucleotide involved in the branch structure and lariat formation. However, in most cases studied to date, the 3' splice site occurs at the first AG downstream of the branch point (Keller, 1984). In intron 7 of the K^k gene the first AG dinucleotide appears just nine nucleotides downstream of the exchanged adenosine (Figure 8). This distance might be too short to allow splicing at this site. For other introns the corresponding number of nucleotides varies between 18 and 37 (Konarska *et al.*, 1985; Reed and Maniatis, 1985; Ruskin *et al.*, 1985). Thus, if this adenosine is used as the branch point, the second AG would presumably be used, and explain the K^k splicing.

The branch point adenosine in the K^k gene has been exchanged for a thymidine in the K^b gene and therefore cannot be involved in the lariat formation. The K^b branch point must be present further upstream. The first reasonable match to the PyTPuAPy consensus is the sequence aGTGATgg with the A-residue 28 nucleotides upstream of the AG splice site (consensus shown in capitals). We believe that this is the branch point in intron 7 of the K^b gene. Interestingly, the K^k sequence contains an identical sequence aGTGATgg 25 nucleotides upstream of its AG splice site. In the K^b gene this latter sequence is part of exon 8. The sequences are shown in boxes in Figure 8.

It is possible to develop a clearer understanding of the splicing of intron 7 of the K^k and K^b genes by studying the corresponding intron in other H-2 class I genes. We have examined the DNA sequences of intron 7 for five murine class I genes and compared them on the basis of the known splicing patterns deduced from the corresponding cDNA sequences. The DNA sequence in the 3' splicing region of intron 7 of the K^d gene (Kvist *et al.*, 1983) is identical to that of the K^b gene (Weiss *et al.*, 1983). In both these genes exon 8 encodes 10 amino acids and is spliced identically (Xin *et al.*, 1982; Lalanne *et al.*, 1983; Reyes *et al.*, 1982). The D^d gene (Sher *et al.*, 1985) has the same branch point adenosine residue as the K^k gene, is spliced in the same way (Brégégère *et al.*, 1981; Sher *et al.*, 1985) and contains only one amino acid in exon 8. The L^d gene is also spliced to give

only one amino acid in exon 8 (P.Kourilsky, personal communication), although it lacks the adenosine residue present nine nucleotides upstream of the splice site in the K^b and K^d genes (Figure 8; Moore *et al.*, 1982).

From these observations we interpret the following. First, the K^k and D^d genes contain the branch site consensus sequence (Figure 7; site 1, the TTGAC sequence). The acceptor site AG, nine nucleotides downstream, is too close for efficient splicing and so the acceptor site 36 nucleotides downstream is used in both genes to give a single amino acid encoded by exon 8. Second, there are two potential branch sites in the K^b and K^d genes (both GTGAT, sites 2 and 3 in Figure 8) of which both genes utilize the branch point at site 2 and the acceptor site 28 nucleotides downstream to give 10 amino acids encoded by exon 8. These two genes cannot use the K^k, D^d branch site (site 1) as this has been inactivated by the T for A substitution. Third, in the L^d gene, sites 1 and 2 cannot be used as the key branch point adenosine residues have been substituted by a thymidine and guanosine residue, respectively. Therefore, the L^d gene must use a different branch site (Figure 8, site 3, the GTGAT sequence) and the AG acceptor site 25 nucleotides downstream. This yields the eighth exon encoding a single amino acid. From this comparative study, we can rank these putative branch sites. All three branch sites are present in the K^k gene but the results from the mutagenesis experiments clearly implicate site 1. Both sites 2 and 3 are present in the K^b and K^d genes, but both these genes have 10 amino acids encoded by exon 8 and therefore must use site 2. The L^d gene can only use one site — site 3. The apparent order of preference therefore must be site 1 followed by site 2 and finally site 3 (Figure 8). Thus, the sequence TGAPy seems crucial for 3' splicing of intron 7. Preferably this sequence should be preceded by a thymidine residue (site 1, TTGAC in the K^k and D^d genes), but also a guanosine residue is functional at this position (sites 2 and 3, GTGAT in the K^b, K^d and L^d genes).

All introns examined so far begin with the dinucleotide GT (Mount, 1982) which is the 5' consensus sequence (Breathnach *et al.*, 1978). It has been shown that the first six nucleotides of introns are conserved (Wieringa *et al.*, 1983) and, for the rabbit β -globin gene, mutation of the GT to AT at the start of the in-

tron prevents normal excision of the intron (Wieringa *et al.*, 1984). Similarly, for the β^0 -thalassemic gene, the GT to AT at position 1 in the second intron causes abnormal splicing of the intron (Treisman *et al.*, 1982). Deletion of five nucleotides (position 2–6) from the first intron of the α 2-globin gene (α -thalassemia) inactivates the normal donor splice site and activates an alternative donor site (Felber *et al.*, 1982). This deletion converts the GT consensus sequence to GC. For a β -thalassemic gene, positions 5 and 6 of the first intron do not inactivate the site, but reduce its activity (Treisman *et al.*, 1983). Thus, the first six nucleotides of introns seem crucial for correct and efficient splicing.

The third nucleotide of intron 7 is different between the K^b and K^k genes. The K^b gene has a G at this position whereas the K^k gene has an A residue. Both genes are correctly spliced indicating that this nucleotide position is of less importance. Also, the 5' mutation in intron 7 of the H-2 K^k gene (m77) does not change the splicing pattern (Figure 7B). A purine nucleotide at this position might, however, be a requirement (Breathnach *et al.*, 1978; Lewin, 1980).

The aim of this study was to explain the profound effect on the splicing of the K^b and K^k genes in terms of the two nucleotide differences in intron 7. We have been able to map this effect to a single nucleotide difference at the 3' end of the intron. Both the H-2 K^b and K^k antigens can function as restricting elements during virus infection (Zinkernagel and Doherty, 1979) and deletion or changes of the cytoplasmic part of class I histocompatibility antigens do not change this ability (Zuniga *et al.*, 1983; Murre *et al.*, 1984). It is therefore impossible at present to assess the eventual cellular consequences of the observed phenomenon and further information must be obtained on the function of the cytoplasmic part of histocompatibility class I antigens.

Materials and methods

Cell culture and DNA transfection of cells

The mouse fibroblast cell line 1T 22-6 (H-2^b) has been used in this study. Cells were grown in Dulbecco's minimal essential medium (DMEM) containing 10% fetal calf serum. Transfection of cells was done by a modification of the method of Wigler *et al.* (1979) and has been described in detail (Arnold *et al.*, 1984a). The neo-phosphotransferase gene (Southern and Berg, 1982) was used as the selectable marker in combination with the antibiotic G-418 (Gibco).

Cell labelling, antisera, immunoprecipitation and SDS-PAGE

Labelling of cells with [³⁵S]methionine and immunoprecipitation were carried out as previously described (Arnold *et al.*, 1984a). The monoclonal antibody against the H-2 K^b was 28-8-6S (Ozato and Sachs, 1981). This antibody was used to precipitate the K^b and the C2 antigens. To identify the K^k antigen, the C4 antigen and the products of the m77 and m80 genes, we have used the antibody H100-27.55 (Lemke *et al.*, 1979). Immunoprecipitation and SDS-PAGE were carried out as described previously (Kvist *et al.*, 1982).

Fluorescence-activated cell sorter (FACS) analysis

To ascertain cell surface expression of the authentic H-2 K^b and K^k antigens, the C2 and C4 antigens and the products of the m77 and m80 genes, transfected cells were analysed by cytofluorography. The same antibodies were used as for the immunoprecipitation. The labelling procedure has been outlined earlier (Burgert and Kvist, 1985). The cells were analysed in a FACS II cytofluorograph.

Construction of hybrid genes and site-directed mutagenesis

The construction of the pC4 gene was made in a similar way as described for the pC2 gene (Arnold *et al.*, 1984a). The *Bg*III/*Sal*I fragment (3' end) of the K^b gene, was combined with the *Sal*I/*Bg*III fragment (5' end) of the K^k gene in pBR322 (see Figure 1B). Correct joining of the fragments was ascertained by determining the DNA sequence around the *Bg*III site and by constructing a partial restriction map of the hybrid gene.

The two 19-mer oligonucleotides (K^k 4606 and K^k 4707) were synthesized manually by the phosphate methodology and purified by h.p.l.c. as described by Connolly *et al.* (1984).

The site-directed mutagenesis strategy is outlined in Figure 5. The template for the mutagenesis was a subclone (pK^k -16-5) of the 3' end of the H-2 K^k gene

which had been derived using the method of Frischauf *et al.* (1980). Plasmid DNA from clone (pK^k -16-5) was isolated from the *E. coli* strain DH-1 (Hanahan, 1983).

Covalently closed circular DNA (pK^k -16-5) was nicked with the restriction endonuclease *Cl*aI by incubating 60 μ g of plasmid DNA with 100 units of *Cl*aI in 500 μ l of 10 mM Tris-HCl, pH 8.0, 10 mM MgCl₂, 10 mM dithiothreitol (DTT) and 75 μ g/ml of ethidium bromide for 60 min at 25°C. About 75% of the DNA was converted to a relaxed circular form (nicked). The reaction was stopped with excess EDTA prior to extraction three times with 10 volumes of water-saturated isobutanol, once with 2 volumes of phenol:chloroform:isoamyl-alcohol (25:24:1) and then twice with 10 volumes of water-saturated diethylether (Wallace *et al.*, 1980). The volume was reduced by extraction with isobutanol prior to ethanol precipitation.

These extensive organic extractions were necessary to remove all traces of ethidium bromide which potentially inhibit the next enzyme – exonuclease III. Although the nicking occurs specifically at the *Cl*aI recognition site, only 50% of the relaxed DNA will be nicked on the desired strand.

The nicked DNA was rendered partially single stranded by digestion with exonuclease III. The conditions recommended by Guo and Wu (1983) were used, i.e. 20 units of exonuclease III/pmol of DNA in 50 mM Tris-HCl, pH 8.0, 90 mM NaCl, 5 mM MgCl₂, 10 mM DTT at 23°C for 90 min. Under these conditions ~1200 nucleotides are removed from each plasmid in a synchronous manner. The reaction was stopped with excess EDTA and the DNA recovered after phenol–chloroform extraction and ethanol precipitation.

The template DNA was then digested with *Bg*III. This digestion linearized any residual supercoiled DNA as well as those plasmids in which the single-stranded region is in the vector sequences rather than in the target area of the gene (see Figure 5). The DNA recovered following phenol–chloroform extraction and ethanol precipitation was used as the target/template for the oligonucleotide-directed mutagenesis.

Phosphorylated oligonucleotide (~200 pmol) was annealed to the target DNA (~5 pmol) by incubating at 55°C for 10 min followed by 10 min on ice in 20 mM Tris-HCl, pH 7.5, 50 mM MgCl₂, 1 mM DTT.

Extension and ligation was effected overnight at 15°C with 5 units of DNA polymerase I (Klenow fragment) and 1 unit of T4 DNA ligase in 20 mM Tris-HCl pH 7.5, 10 mM MgCl₂, 10 mM DTT, 0.5 mM each of dNTPs and 0.5 mM ATP. Aliquots of this reaction mixture were used to transform the *E. coli* strain 71/18 (Dente *et al.*, 1983).

Colonies (~250 per experiment) were picked in duplicate onto nitrocellulose filters, grown for 4–6 h at 37°C on ampicillin plates and then amplified overnight on chloramphenicol plates (the replica filters were not amplified, but were stored at 4°C). Colonies of bacteria with plasmids carrying the H-2 K^k and H-2 K^b genes were used as negative and positive controls, respectively. The filters were placed for 5 min each on 3MM filter paper soaked in the following solutions: (i) 1% SDS, (ii) 0.5 M NaOH, 1.5 M NaCl, (iii) 1 M NH₄Ac, 0.02 M NaOH, (iv) as for (iii), (v) 2 × SSPE, 0.1% SDS and then immersed for 45 s in chloroform, rolled between sheets of 3MM paper and then baked at 80°C for 2 h.

Filters were soaked for 15 min in 6 × NET (0.9 M NaCl, 90 mM Tris-HCl, pH 8, 6 mM EDTA), 0.1% SDS at room temperature. Pre-hybridization was carried out for 3 h at 65°C in 6 × NET, 10 × Denhardt's, 0.1% SDS (sonicated denatured salmon sperm DNA was added after 1 h). Hybridization was performed in 6 × NET, 10 × Denhardt's, 0.2% SDS, plus labelled oligonucleotide, at 45°C overnight (2 × 10⁶ c.p.m./ml, 1 × 10⁸ c.p.m./ μ g).

The following day the filters were washed as follows: (i) three changes of 6 × SSC in 30 min at room temperature, (ii) 6 × SSC at 45°C (10 min for K^k 4606; 15 min for K^k 4707), (iii) 6 × SSC at 48°C (5 min for K^k 4606; 15 min for K^k 4707), (iv) two changes of 6 × SSC at 54°C in 30 min for K^k 4707 only. After each wash the filters were placed with X-ray film plus two intensifying screens at –70°C for 1 h.

Rapid plasmid-purified DNA was isolated from putative mutant colonies and used to transform *E. coli* 71/18 bacteria which were then picked (only 16/transformation) in duplicate and screened as described above.

Plasmid DNA was isolated from four independent colonies for each mutant shown to be positive in the second round of screening and sequenced by the method of Maxam and Gilbert (1980).

S1 nuclease mapping

We have used a subclone of the H-2 K^k gene (pK^k -16-5) to construct the probe for the S1 nuclease mapping (Figure 7A). This subclone originates from a subcloning deletion in the 5'–3' deletion of the H-2 K^k gene (Frischauf *et al.*, 1980; Arnold *et al.*, 1984a). This clone contains part of exon 5 (the exon encoding the transmembrane segment), exons 6, 7 and 8, as well as all the intervening sequences through to the 3'-untranslated region of exon 8. The clone starts at the 5' end with a *Cl*aI site and carries 23 nucleotides upstream the *Eco*RI site of pBR322. In the untranslated region of exon 8, a *Sau*96I site is present 164 nucleotides downstream of the termination codon. DNA of clone pK^k -16-5 was

digested with the *Sau96I* enzyme and the ends were made blunt with the Klenow DNA polymerase in the presence of dNTPs. A second digestion step with *EcoRI* cuts the DNA only once (in the pBR322 sequence) and yields a 776-bp fragment which was inserted into the *EcoRI/HincII* sites of M13mp8. This M13 clone was termed 16-5mp8 and has the orientation such that the sense strand of the DNA is present in phages secreted from transfected bacteria.

By using 16-5mp8 single-stranded DNA (ssDNA) as a template, a uniformly ^{32}P -labeled anti-sense strand was synthesized. 1 pmol of 16-5mp8 ssDNA was annealed with a 10-fold excess of the pentadecamer primer in 20 μl of 100 mM NaCl, 20 mM Tris-HCl, pH 7.9, 20 mM MgCl_2 and 20 mM β -mercaptoethanol. The solution was heated for 2 min at 90°C, incubated at 65°C for an additional 15 min and then cooled on ice. To the annealing reaction was added 20 μl of a solution containing 200 μM each of dATP, dGTP, and dTTP, 20 μM dCTP, 2 μM [α - ^{32}P]dCTP (800 Ci/mmol) and 5 units of the Klenow DNA polymerase. Polymerization was allowed to proceed at 37°C for 30 min and was terminated by heating for 15 min at 70°C.

The mixture was restricted with the *EcoRI* enzyme and then diluted 2-fold by the addition of 60% dimethylsulphoxide, 2 mM EDTA, 0.1% xylene cyanol and bromophenol blue. After heating for 2 min at 90°C, cooled on ice, the fragments were separated on a 4% polyacrylamide gel at 150 V for 16 h. The gel was exposed under X-ray film to identify the 776-nucleotide fragment. This band was cut out of the gel, crushed and eluted at 37°C for 4 h in 1.5 ml of Maxam and Gilbert elution buffer (Maxam and Gilbert, 1980). The aqueous phase was collected by passing the mixture through a syringe plugged with glass wool. The probe was phenol extracted once and concentrated by ethanol precipitation.

mRNA was isolated from DNA-transfected cells by using Vanadyl complex for protection against nucleases and subsequently extracted with phenol (Maniatis *et al.*, 1982). 4 μg of mRNA from each cell clone was hybridized with 8×10^4 c.p.m. of 16-5mp8 probe, in the presence of 80% formamide, 40 mM Pipes pH 6.4, 1 mM EDTA and 0.4 M NaCl. Samples were heated to 85°C for 5 min and then incubated at 52°C for 3 h. After the incubation period, 200 μl of ice-cold S1 buffer (280 mM NaCl, 50 mM sodium acetate pH 4.6, and 4.5 mM ZnSO_4) and 10 units/ μl of S1 nuclease were added. Incubation was at 45°C for 30 min (Maniatis *et al.*, 1982). The samples were then ethanol precipitated and analyzed on a 6% polyacrylamide DNA sequencing gel. In parallel, the clone 16-5mp8 which had undergone dideoxy sequencing reactions was analysed (Sanger *et al.*, 1977). After separation of the fragments, the gel was soaked for 10 min in 10% acetic acid, dried on the glass plate and the bands were visualized by exposure to X-ray film.

Enzymes and reagents

Restriction endonucleases and other DNA-modifying enzymes were from Boehringer Mannheim, New England Biolabs and PL-Pharmacia Biochemicals. Geneticin (G-418) was purchased from Gibco. Protein A-Sepharose was from Pharmacia Fine Chemicals, Uppsala, Sweden. [^{35}S]Methionine, [^{32}P]dNTPs and [γ - ^{32}P]ATP were from Amersham, UK. The M13 pentadecamer sequencing primer was from New England Biolabs.

Acknowledgements

We would like to thank Drs B. Arnold, H.-G. Burgert and D. Cutler for useful discussions and Dr P. Kourilsky for communicating unpublished results. We thank Dr B. Connolly for his guidance on the oligonucleotide synthesis. For critical reading of the manuscript we thank Drs W. Jefferies and H.-G. Burgert. For technical assistance we thank C. Raynoschek, and for the FACS analysis P. Zaeck. Excellent help in preparing the manuscript was provided by C. Droz and C. Ravussin. A.L.A. was supported by the Agricultural and Food Research Council, London. This work was supported by a grant to S. Kvist from the Swiss National Science Foundation.

References

- Arnold, B., Burgert, H.-G., Hamann, U., Hämmerling, G., Kees, U. and Kvist, S. (1984a) *Cell*, **38**, 79–87.
- Arnold, B., Burgert, H.-G., Archibald, A.L. and Kvist, S. (1984b) *Nucleic Acids Res.*, **12**, 9473–9487.
- Breathnach, R., Benoist, C., O'Hare, K., Gannon, F. and Chambon, P. (1978) *Proc. Natl. Acad. Sci. USA*, **75**, 4853–4857.
- Brégère, F., Abastado, J.P., Kvist, S., Rask, L., Lalanne, J.L., Garoff, H., Cami, B., Wiman, K., Larhammar, D., Peterson, P.A., Gachelin, G., Kourilsky, P. and Dobberstein, B. (1981) *Nature*, **292**, 78–81.
- Burgert, H.-G. and Kvist, S. (1985) *Cell*, **41**, 987–997.
- Connolly, B.A., Potter, B.V.L., Eckstein, F., Pingoud, A. and Grotjahn, L. (1984) *Biochemistry*, **23**, 3443–3453.
- Dente, L., Cesareni, G. and Cortese, R. (1983) *Nucleic Acids Res.*, **11**, 1645–1655.
- Domdey, H., Apostol, B., Lin, R.-J., Newman, A., Brody, E. and Abelson, J. (1984) *Cell*, **39**, 611–621.
- Felber, B.K., Orkin, S.H. and Hamer, D.H. (1982) *Cell*, **29**, 895–902.
- Freundewey, D. and Keller, W. (1985) *Cell*, **42**, 355–367.
- Frischauf, A.M., Garoff, H. and Lehrach, H. (1980) *Nucleic Acids Res.*, **8**, 5541–5549.
- Grabowski, P.J., Padgett, R.A. and Sharp, P.A. (1984) *Cell*, **37**, 415–427.
- Green, M.R., Maniatis, T. and Melton, D.A. (1983) *Cell*, **32**, 681–694.
- Guo, L.-H. and Wu, R. (1983) *Methods Enzymol.*, **100**, 60–96.
- Hanahan, D. (1983) *J. Mol. Biol.*, **166**, 557–580.
- Hernandez, N. and Keller, W. (1983) *Cell*, **35**, 89–99.
- Keller, E.B. and Noon, W.A. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 7417–7420.
- Keller, W. (1984) *Cell*, **39**, 423–425.
- Konarska, M.M., Grabowski, P.J., Padgett, R.A. and Sharp, P.A. (1985) *Nature*, **313**, 552–557.
- Krainer, A.R., Maniatis, T., Ruskin, B. and Green, M.R. (1984) *Cell*, **36**, 993–1005.
- Kress, M., Glaros, D., Khoury, G. and Jay, G. (1983) *Nature*, **306**, 602–604.
- Kvist, S., Wiman, K., Claesson, L., Peterson, P.A. and Dobberstein, B. (1982) *Cell*, **29**, 61–69.
- Kvist, S., Roberts, L. and Dobberstein, B. (1983) *EMBO J.*, **2**, 245–254.
- Lalanne, J.L., Delarbre, C., Gachelin, G. and Kourilsky, P. (1983) *Nucleic Acids Res.*, **11**, 1567–1577.
- Langford, C.J., Klinz, F.-J., Donath, C. and Gallwitz, D. (1984) *Cell*, **36**, 645–653.
- Lemke, H., Hämmerling, G.J. and Hämmerling, U. (1979) *Immunol. Rev.*, **47**, 175–206.
- Lewin, B. (1980) *Cell*, **22**, 324–326.
- Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) *Molecular Cloning. A Laboratory Manual*. Cold Spring Harbor Laboratory Press, NY.
- Maxam, A.M. and Gilbert, W. (1980) *Methods Enzymol.*, **65**, 499–560.
- Moore, K.W., Sher, B.T., Sun, Y.H., Eakle, K.A. and Hood, L. (1982) *Science*, **215**, 679–682.
- Mount, S.M. (1982) *Nucleic Acids Res.*, **10**, 459–472.
- Murre, C., Reiss, C.S., Bernabeu, C., Chen, L.B., Burakoff, S.J. and Seidman, J.G. (1984) *Nature*, **307**, 432–436.
- Ozato, K. and Sachs, D.H. (1981) *J. Immunol.*, **126**, 317–321.
- Padgett, R.A., Hardy, S.F. and Sharp, P.A. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 5230–5234.
- Padgett, R.A., Konarska, M.M., Grabowski, P.J., Hardy, S.F. and Sharp, P.A. (1984) *Science*, **225**, 898–903.
- Pikielny, C.W., Teem, J.L. and Rosbash, M. (1983) *Cell*, **34**, 395–403.
- Reed, R. and Maniatis, T. (1985) *Cell*, **41**, 95–105.
- Reyes, A.A., Schödl, M., Itakura, K. and Wallace, R.B. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 3270–3274.
- Rodriguez, J.R., Pikielny, C.W. and Rosbash, M. (1984) *Cell*, **39**, 603–610.
- Ruskin, B. and Green, M.R. (1985) *Nature*, **317**, 732–734.
- Ruskin, B., Krainer, A.R., Maniatis, T. and Green, M.R. (1984) *Cell*, **38**, 317–331.
- Ruskin, B., Greene, J.M. and Green, M.R. (1985) *Cell*, **41**, 833–844.
- Sanger, F., Nicklen, S. and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
- Sharp, P.A. (1985) *Cell*, **42**, 397–400.
- Sher, B.T., Nairn, R., Coligan, J.E. and Hood, L.E. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 1175–1179.
- Southern, P.J. and Berg, P. (1982) *J. Mol. Appl. Genet.*, **1**, 327–341.
- Steinmetz, M. and Hood, L. (1983) *Science*, **222**, 727–733.
- Treisman, R., Proudfoot, N.J., Shander, M. and Maniatis, T. (1982) *Cell*, **29**, 903–911.
- Treisman, R., Orkin, S.H. and Maniatis, T. (1983) *Nature*, **302**, 591–596.
- Wallace, J.C. and Edmonds, M. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 950–954.
- Wallace, R.B., Johnson, P.F., Tanaka, S., Schödl, M., Itakura, K. and Abelson, J. (1980) *Science*, **209**, 1396–1400.
- Wallace, R.B., Schödl, M., Johnson, M.J., Dembek, P. and Itakura, K. (1981) *Nucleic Acids Res.*, **9**, 3647–3656.
- Weiss, E., Golden, L., Zakut, R., Mellor, A., Fahrner, K., Kvist, S. and Flavell, R.A. (1983) *EMBO J.*, **2**, 453–462.
- Weissmann, C. (1984) *Nature*, **311**, 103–104.
- Wieringa, B., Meyer, F., Reiser, J. and Weissmann, C. (1983) *Nature*, **301**, 38–43.
- Wieringa, B., Hofer, E. and Weissmann, C. (1984) *Cell*, **37**, 915–925.
- Wigler, M., Pellicer, A., Silverstein, S., Axel, R., Urlaub, G. and Chasin, L. (1979) *Proc. Natl. Acad. Sci. USA*, **76**, 1373–1376.
- Xin, J.-H., Kvist, S. and Dobberstein, B. (1982) *EMBO J.*, **1**, 467–471.
- Zeitlin, S. and Efstratiadis, A. (1984) *Cell*, **39**, 589–602.
- Zinkernagel, R.M. and Doherty, P.C. (1979) *Adv. Immunol.*, **27**, 51–177.
- Zuniga, M.C., Malissen, B., McMillan, M., Brayton, P.R., Clark, S.S., Forman, J. and Hood, L. (1983) *Cell*, **34**, 535–544.

Received on 15 January 1986